Results From The UKQCD Parallel Tempering Project

B. Joó[a], *UKQCD Collaboration*

[a]Department of Physics and Astronomy, University of Edinburgh,
James Clerk Maxwell Building, The Kings Buildings, Mayfield Road, Edinburgh EH9 3JZ, Scotland, UK

We present results from our study of the Parallel Tempering algorithm. We examine the swapping acceptance rate of a twin subensemble PT system. We use action matching technology in an attempt to maximise the swap acceptance rate. We model the autocorrelation times within Parallel Tempering ensembles in terms of autocorrelation times from Hybrid Monte Carlo. We present estimates for the autocorrelation times of the plaquette operator.

1. INTRODUCTION

Dynamical fermion simulations are still very demanding computationally. While a modern supercomputer is able to produce $O(100)$ quenched gauge configurations in an afternoon, months are needed to generate a similar number of dynamical configurations on the same machine.

Tempering algorithms have been successful in the past in reducing autocorrelation times in difficult spin glass simulations such as the random field Ising model [1]. It was of interest to see whether parallel tempering (PT) could achieve the same success for lattice QCD.

We present only our main results in this publication. Full details of our work can be found in [2]. A lot of the background information is covered in [3–5].

In section 2 we outline the PT algorithm. In section 3 we present the predictions of a simple model of autocorrelations in a twin subensemble PT system. Our simulation parameters are outlined in 4. Results from the simulations are presented in section 5. Our summary and conclusions are in section 6

2. THE ALGORITHM

Parallel tempering consists of running several simulations, hereafter referred to as *subensembles*, concurrently. Each subensemble has its own Hamiltonian $\mathcal{H}_i$, parameter set and phase space. The overall PT state then is the set $\{s_i | i = 1...N\}$ where s_i is the state of subensemble i and N is the number of subensembles. The PT phase space is the direct product of the individual phase spaces.

The goal of PT is to construct a Markov Process which will converge to the equilibrium probability distribution

$$P_{eq} = \prod_i \frac{1}{Z_i} e^{-S_i}, \tag{1}$$

where the terms on the right hand side are the equilibrium distributions of the individual subensembles at their given parameter sets.

One defines two kinds of Markov transitions:

- transitions within subensembles
- transitions between subensembles

Transitions within subensembles are carried out using HMC. Transitions between subensembles involve a proposal to swap the current fields in a subensemble i with those in subensemble j. The swap proposal is accepted with the Metropolis acceptance probability

$$\mathcal{P}_s(i,j) = \min\left(1, e^{-\Delta\mathcal{H}}\right), \tag{2}$$

where

$$\Delta\mathcal{H} = \{\mathcal{H}_j(a) + \mathcal{H}_i(b)\} - \{\mathcal{H}_i(a) + \mathcal{H}_j(b)\}, \tag{3}$$

which satisfies detailed balance with respect to P_{eq} by construction. The resulting overall Markov Process is connected and satisfies detailed balance with respect to the required equilibrium.

2.1. Swap Acceptance

The acceptance rate of swap proposals determines any reduction in autocorrelation times over the usual HMC ones. It is also related to a distance in parameter space via the formalism of action matching technology [3]. It can be shown from detailed balance considerations that the swap acceptance rate is given by

$$\langle A \rangle = \mathrm{erfc}\left(\frac{1}{2}\sqrt{\langle \Delta\mathcal{H} \rangle}\right). \tag{4}$$

3. AUTOCORRELATIONS

Consider a twin subensemble system where both subensembles have the same HMC autocorrelation function $C_H(t)$. This assumption is justifiable if the HMC autocorrelation times vary slowly over the region of parameter space where one intends to carry out PT simulations. Table 1 shows the integrated autocorrelation times (τ_{H}) of the plaquette for two HMC simulations separated by such a distance in parameter space. It can be seen that the integrated autocorrelation times are equal within errors.

Let us assume that individual swap probabilities may be replaced by the average swap probability $\langle A \rangle$. Furthermore we are interested only in an even number of swap attempts. Only after an even number of successful swaps can a configuration end up in its original subensemble. We assume no cross correlations between subensembles.

It can be shown [2] that the resulting PT autocorrelation function in each subensemble is:

$$C_{PT}(t) = \frac{1}{2}\left\{1 + (1 - 2\langle A \rangle)^t\right\} C_H(t) \tag{5}$$

and assuming an exponentially decaying HMC autocorrelation function, the ratio of the HMC integrated autocorrelation time to its PT counterpart (τ_{PT}) is given by

$$\frac{\tau_{PT}}{\tau_H} = \frac{\langle A \rangle(\tau_H - 1) + 1}{1 + 2\langle A \rangle \tau_H}. \tag{6}$$

The above ratio is bounded above by 1 and tends to $\frac{1}{2}$ as τ_H increases for a fixed $\langle A \rangle$. Hence this model predicts that the expected gain from PT over HMC is a factor of 2 in each subensemble for roughly twice the work. Models of systems with different HMC autocorrelation times are under investigation.

Table 1
HMC integrated autocorrelation times for the plaquette

β	κ	τ_{H}
5.2	.1335	18(8)
5.232	.1335	20(6)

4. SIMULATION

Our simulations used the GHMC [6] code developed by the UKQCD collaboration for HMC transitions with extra logic to carry out the swapping. The simulation parameters to be tuned were the inverse gauge coupling β, the fermion hopping parameter κ and the clover coefficient c.

We carried out five twin subensemble simulations: $S1, S2, S3, S4$ and $S5$. The first subensembles of each one had parameters $(\beta_1, c_1, \kappa_1) = (5.2, 2.0171, 0.1330)$. The parameters for the second subensembles are shown in table 2. The simulation parameters for $S1$, $S2$ and $S3$ were tuned using action matching [3]. For $S4$ and $S5$ only κ was varied. The lattice size used was $8^3 \times 16$. A reference HMC run at the parameters of the first ensembles was also carried out.

5. RESULTS

Table 3 summarises the results of our simulations. Looking at columns 2, 3 and 4 it can be seen that $\langle \Delta\mathcal{H} \rangle$ is usually greater than one and that $\langle A \rangle$ drops rapidly as the magnitude of

Table 2
Simulation parameters for second ensembles

Simulation	(β_2, c_2, κ_2)
$S1$	$(5.2060, 2.01002, 0.13280)$
$S2$	$(5.2105, 2.00471, 0.13265)$
$S3$	$(5.2150, 1.99940, 0.13250)$
$S4$	$(5.2, 2.0171, 0.13280)$
$S5$	$(5.2, 2.0171, 0.13265)$

Table 3
Simulation Results

Simulation	$\Delta\kappa(\times 10^{-4})$	$\langle\Delta\mathcal{H}\rangle$	$\langle A\rangle$	$\tau_{\rm int}$	$\frac{\tau_{\rm PT}}{\tau_{\rm H}}$
HMC	-	-	-	26(6)	1
$S1$	-2.0	1.23(2)	0.43(1)	12(3)	0.5(2)
$S2$	-3.5	3.76(4)	0.17(1)	19(4)	0.7(2)
$S3$	-5.0	7.64(6)	0.051(2)	24(6)	0.9(3)
$S4$	-2.0	0.91(4)	0.49(1)	9(4)	0.3(2)
$S5$	-3.5	2.29(7)	0.26(2)	18(10)	0.7(4)

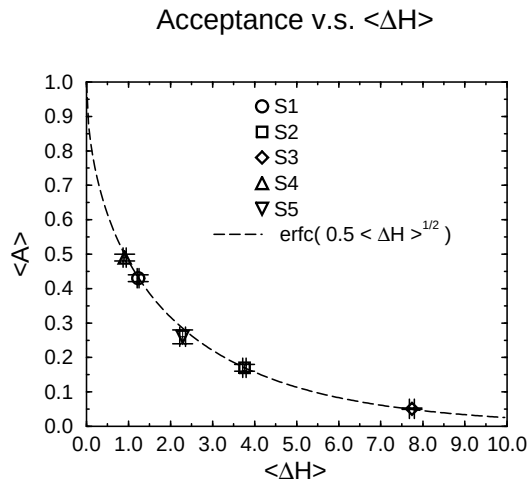

Figure 1. Swap Acceptance Rates

$\Delta\kappa = \kappa_2 - \kappa_1$ is increased. Figure 1 shows the acceptance rate $\langle A\rangle$ as a function of $\langle\Delta\mathcal{H}\rangle$. The dashed line is the graph of (4). It can be seen that the measurements and the predictions of the acceptance rate agree very well.

We note that the simulations with parameters given by action matching technology ($S1 - S3$) have a lower acceptance rate than the others. This is because the fluctuations in $\Delta\mathcal{H}$ are larger using our pseudofermionic Hamiltonian, than in the action used to perform the matching. This issue is discussed more fully in [2].

Column 5 in table 3 shows the integrated autocorrelation times of the plaquette for our simulations. Column 6 gives the corresponding ratios of PT to HMC autocorrelation times. These ratios are consistent with the predictions of the model described in section 3.

6. CONCLUSIONS

We find that the acceptance rate of the PT algorithm drops very rapidly with $\Delta\kappa$. This problem is expected to get worse on larger lattices as $\Delta\mathcal{H}$, an extensive quantity, will have larger fluctuations. Over the range of available $\Delta\kappa$ the HMC autocorrelation times are equal within errors and the predictions of our autocorrelation model apply. We estimate that connecting fast and slow decorrelating regions of parameter space would need a very large number of subensembles making PT impractical to use for lattice QCD with currently available computer technology.

7. ACKNOWLEDGEMENTS

We gratefully acknowledge funding PPARC under grant number GR/L22744 and EPSRC under grant number R/K41663.

REFERENCES

1. E. Marinari, G. Parisi, J. Ruiz-Lorenzo, Spin Glasses and Random Fields, *cond-mat/9701016*
2. B. Joó, A. C. Irving, J. C. Sexton, B. J. Pendleton, S. M. Pickles, Z. Scroczynski, UKQCD Collaboration, in preparation.
3. A. C. Irving, J. C. Sexton, E. Cahill, J. Garden, B. Joó, S. M. Pickles, Z. Sroczynski, *hep-lat/9807015*
4. G. Boyd, Nucl. Phys.(Proc. Suppl.) 60A (1998) 341-344
5. E. Marinari, *cond-mat/9612010*
6. Z. Sroczynski, S. M. Pickles and S. P. Booth, UKQCD Collaboration, Nucl. Phys. (Proc. Suppl.) 63A-C (1998) 949-951.